RESEARCH ARTICLE

Regional models do not outperform continental models for invasive species

Catherine S. Jarnevich[1], Helen R. Sofaer[2], Pairsa Belamaric[3,4], Peder Engelstad[5]

1 *U.S. Geological Survey, Fort Collins Science Center, Fort Collins, CO, USA* **2** *U.S. Geological Survey, Pacific Island Ecosystems Research Center, Hawaii National Park, HI, USA* **3** *Student Contractor to U.S. Geological Survey, Fort Collins Science Center, Fort Collins, CO, USA* **4** *University of Wisconsin, Madison, WI, USA* **5** *Natural Resource Ecology Laboratory, Colorado State University, Fort Collins CO, USA*

Corresponding author: Catherine S. Jarnevich (jarnevichc@usgs.gov)

Abstract

Aim: Species distribution models can guide invasive species prevention and management by characterizing invasion risk across space. However, extrapolation and transferability issues pose challenges for developing useful models for invasive species. Previous work has emphasized the importance of including all available occurrences in model estimation, but managers attuned to local processes may be skeptical of models based on a broad spatial extent if they suspect the captured responses reflect those of other regions where data are more numerous. We asked whether species distribution models for invasive plants performed better when developed at national versus regional extents.

Location: Continental United States.

Methods: We developed ensembles of species distribution models trained nationally, on sagebrush habitat, or on sagebrush habitat within three ecoregions (Great Basin, eastern sagebrush, and Great Plains) for nine invasive plants of interest for early detection and rapid response at local or regional scales. We compared the performance of national versus regional models using spatially independent withheld test data from each of the three ecoregions.

Results: We found that models trained using a national spatial extent tended to perform better than regionally trained models. Regional models did not outperform national ones even when considerable occurrence data were available for model estimation within the focal region. Information was often unavailable to fit informative regional models precisely in those areas of greatest interest for early detection and rapid response.

Main conclusions: Habitat suitability models for invasive plant species trained at a continental extent can reduce extrapolation while maximizing information on species' responses to environmental variation. Standard modeling methods can capture spatially varying limiting factors, while regional or hierarchical models may only be advantageous when populations differ in their responses to environmental conditions, a condition expected to be relatively rare at the expanding boundaries of invasive species' distributions.

Keywords
Early detection and rapid response, invasion risk, model transferability, species distribution models

Introduction

Organisms' responses to environmental variation underlie patterns of distribution and abundance and are the basis for correlative statistical tools such as species distribution models (SDMs; Franklin 2010). Among the challenges of such models are that 1) the relationships between environmental conditions and organismal response can vary over space and time, and 2) outcomes under new conditions are difficult to predict (Sequeira et al. 2018; Yates et al. 2018). These twin challenges, transferability and extrapolation, can point to opposing solutions when interest is in predicting habitat suitability in a region beyond the core of a species' range (Werkowska et al. 2017; Sequeira et al. 2018). Transferability challenges could favor limiting both estimation and prediction to within the region of interest when such data are available and where responses to environmental conditions are thought to be distinctive (Barbet-Massin et al. 2018). However, while a regional approach may capture key limiting factors, it excludes the full range of environmental conditions under which data are available and hence can lead to unnecessary extrapolation and errors in estimated suitability (e.g., Fitzpatrick et al. 2007; Broennimann and Guisan 2008).

Predicting suitability for invasive species exemplifies challenges with both transferability and extrapolation (Elith et al. 2010). Wherever invasive species are still spreading, correlative models can conflate this lag in time (i.e., lack of equilibrium) with a lack of suitability. A common recommendation is to develop the most inclusive view of invasion risk by estimating models based on both the native and invaded ranges to capture the species' complete environmental associations and minimize extrapolation (Fitzpatrick et al. 2007; Broennimann and Guisan 2008). However, where populations are differentiated in their responses to environmental conditions, surveillance and management may be more effectively guided by locally or regionally tuned approaches because of poor model transferability (e.g., Connor et al. 2019; Collart et al. 2021). Studies focused on native species have found modeling intraspecific subsets of the data based on genetic or regional groupings improved distribution model predictions (Chardon et al. 2020). Further, regional models performed better at predicting distributions within the margins of species ranges, where different environmental predictors were most important (Vale et al. 2014; Connor et al. 2019). Marginal or poorly sampled populations may also contribute little to model estimation if training data are

heavily dominated by a better sampled portion of a species' range (Pearman et al. 2010; Hällfors et al. 2016); conversely, limiting model estimation to sparse data in a subset of the range may lead to low model quality. Given the potential for population differentiation within species' invaded ranges (e.g., Colautti and Barrett 2013), models of invasive species' distributions may face important trade-offs between inclusivity versus regional applicability, as well as practical data limitations in newly invaded areas. Meta-analyses of studies that trained models using native range only, invaded range only or global range did not find that global models perform better than models generated in the range of interest, and indicate that superiority of global model performance could be a statistical artefact because test data are not independent (Liu et al. 2020b).

Early detection and rapid response (**EDRR**) activities aim to prevent establishment, spread, and impact through surveillance and rapid management action, and can minimize invasions in new regions (Reaser et al. 2020). Sagebrush (*Artemesia* sp.) habitats in the western United States (U.S.) provide habitat for many wildlife species and face multiple stressors including invasive species, altered fire regimes, climate change, and energy development (Davies et al. 2011; Coates et al. 2016). Crist et al. (2019) have developed a list of invasive plants that have no, patchy, or limited presence on a state-by-state basis within sagebrush habitat. Their approach emphasized the potential for ongoing spread and geographic differences in invasion status, as species that are well established within one state may still warrant EDRR elsewhere. For these regional 'EDRR species', species distribution models can guide surveillance by identifying areas with high invasion risk (Brooks and Klinger 2009). However, one concern we have heard from within the management community is that models trained with a broad geographic extent could miss regionally and locally relevant limiting factors if important signals were swamped by other portions of the range.

For a set of nine species recognized as EDRR targets within sagebrush habitats (Crist et al. 2019), we characterized each species' relationship to sagebrush communities to understand habitat associations and degree of sage specialization. We then compared the performance of national (here used to refer to the conterminous U.S.) versus regional species distribution models. We compared regional models to national, instead of global, models because of the availability of a wider breadth of predictors within the conterminous U.S., including higher quality data than are available globally (e.g., for soils), and finer spatial resolution of predictors focused on the U.S. compared to global versions. Appropriate methods to account for sampling biases are also likely to differ between a native range, where a species is likely closer to equilibrium, and a novel range, which complicates background selection when pooling records from native and invaded ranges (Elith et al. 2010; Jarnevich et al. 2017). In addition, all species in question have been in the U.S. since at least 1957 (based on earliest occurrence records; GBIF.org 2022), giving them time to potentially develop local adaptations and providing numerous occurrence points for model estimation (Liu et al. 2020a; b). Thus, we fit species distribution models for each species across the U.S., from all sagebrush within the U.S., and separately within sagebrush habitats in each of three ecoregions (Great Basin, eastern sage, Great Plains). Models trained on sagebrush only

were fit to allow for any response curves specific to sagebrush habitats (within which models were also tested). We evaluated model performance using withheld spatially independent validation data within each region. We hypothesized that given sufficient data and variation in environmental responses, a regional model evaluated with test data from within the region could outperform a national model. Our results evaluate whether national models can sufficiently capture invasion risk across ecoregions, or whether estimation of models for each region improves the credibility of the outputs for on-the-ground management.

Methods

Study area

We used a combination of level 2 and 3 EPA ecoregion designations (U.S. Environmental Protection Agency 2013) to create three regional study areas: the Great Basin (regions 10.1.8/3/5), eastern sage (region 6.2), and Great Plains (regions 9.2, 9.3, and 9.4) regions. Within each regional boundary, we further restricted the study area for each region to sagebrush habitat, defined by 30 m² cells of greater than 0% sagebrush presence as designated by the National Land Cover Data set (**NLCD**) shrubland sagebrush rangeland fractional component product (Xian et al. 2015).

We created a spatial split of the occurrence data for model validation, as random splits typically underestimate prediction error (Roberts et al. 2017; Fourcade et al. 2018). Within each of the three regions we designated a central longitudinal test strip that contained 10% of the sagebrush cells within the region (Fig. 1; Suppl. material 1: Table S1). Occurrence data points within sage habitat inside these test strips were withheld from model fitting and used to evaluate model performance. In addition to these three regional model estimation extents, we considered two larger spatial extents: the continental U.S., and all sagebrush habitat within the continental U.S., defined as above based on Xian et al. (2015; hereafter "all sage").

Study species

We selected nine plants from a list of invasive species for EDRR activities within states of the eastern sage region (Crist et al. 2019) that had at least 50 training occurrence records and 30 test strip records in at least one study region. These included seven forb/herbs as defined by the USDA PLANTS database (USDA NRCS 2019): *Centaurea diffusa* Lam. [diffuse knapweed], *C. solstitialis* L. [yellow star-thistle], *C. virgata* Lam. [squarrose knapweed], *Chondrilla juncea* L. [rush skeletonweed], *Halogeton glomeratus* (M. Bieb.) C.A. Mey. [saltlover, halogeton], *Rhaponticum repens* (L.) Hidalgo Syn *Acroptilon repens* [Russian knapweed] and *Salvia aethiopis* L. [Mediterranean sage]. We also included two winter annual grass species: *Taeniatherum caput-medusae* (L.) Nevski [medusahead rye] and *Ventenata dubia* (Leers) Coss. [ventenata, North Africa grass].
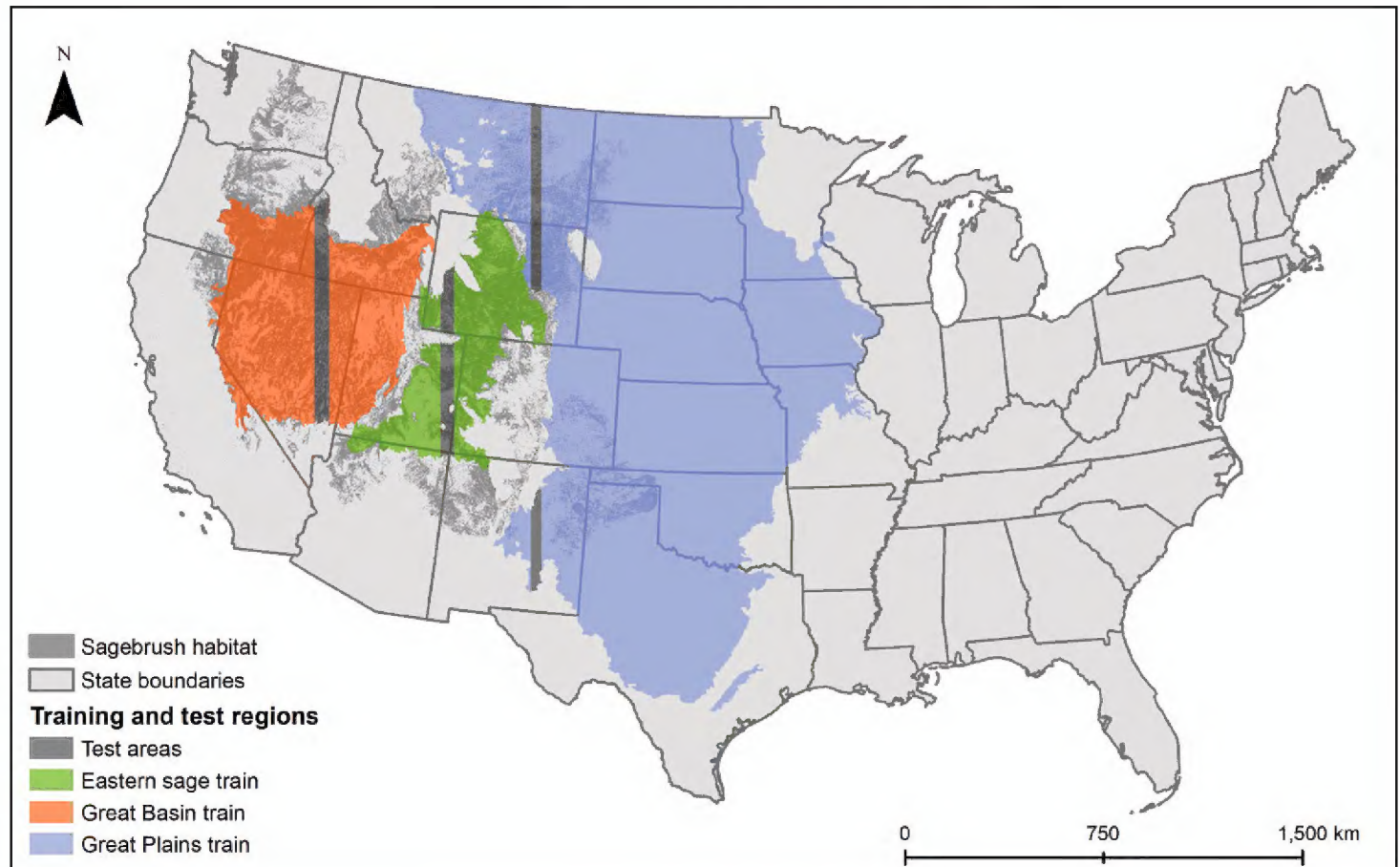
Figure 1. We compared five geographic extents for model estimation while holding validation data constant (occurrence points within dark grey vertical shaded areas). Two geographic training extents were continental and three were regional, and we fit an ensemble of distribution models to the occurrence points for each species within each estimation extent. These extents for model estimation were: 1) the continental United States; 2) all sagebrush habitat within the continental U.S. (gray shading within the western U.S.); 3) sagebrush within eastern sage; 4) sagebrush within the Great Basin; and 5) sagebrush within the Great Plains. Within each of the three regions (shown via colored polygons), we created a test strip (vertical shaded areas) centered on sagebrush habitats, and withheld occurrence points for model performance comparisons. We asked whether a regional or continental training extent yielded higher performance within these test strips, as measured by the Boyce index values.

We aggregated occurrence data from existing data sets following Young et al. (2020), including Global Biodiversity Information Facility (GBIF; GBIF.org 2021), Biodiversity Information Serving Our Nation (BISON; individual data sources and all data used in this study are available as a USGS data release Jarnevich et al. 2022), the Early Detection and Distribution Mapping System (EDDMapS) [all accessed 1 Dec 2020], the Bureau of Land Management's (BLM) Assessment, Inventory and Management database (obtained June 2020) and the BLM and National Park Service's National Invasive Species Information Management System (**NISIMS**) (Suppl. material 1: Table S2 [provides more details in an assessment rubric]). All known synonyms and U.S. Department of Agriculture (**USDA**) Plants Database acronyms were collected (excluding subspecies, variants, and hybrids) using the Integrated Taxonomic Information System (**ITIS**; www.itis.gov) as an authoritative taxonomy in the R library 'taxize' (Chamberlain and Szocs 2013; Chamberlain et al. 2020). We filtered observations by coordinate uncertainty (≤ 30 m), observation type (observation or

specimen only), and observation date (1980 to present [2020]) to match the time frame of predictors and remove older records which typically have poor geographic accuracy. We removed any records with coordinates corresponding to state or country centroids or other easily identifiable geographic and taxonomic errors. We also checked the entire dataset for duplicate records and confirmed that occurrence locations generally aligned with reported distributions via USDA Plants Database (USDA NRCS 2019). We followed these same methods to obtain location data for species identified as non-native by USDA Plants Database to use as background points to control for sampling biases, as described below. We required 50 occurrence records in a study area to fit a model, and 30 records in a test strip to evaluate models for that strip (Suppl. material 1: Table S3, Fig. S1).

Predictors

We began with a national library of 49 predictors representing climate (water deficit, actual evapotranspiration, precipitation, and temperature average from available years 1981–2018 [see Suppl. material 1: Table S4]), human disturbances, soils, water presence / recurrence, fire history, and land cover created by Young et al. (2020) using the Albers equal area projection with a 90 m^2 resolution and modified by Engelstad et al. (2022) (Suppl. material 1: Table S4). This list includes predictors thought to be important for determining the distribution of different types of plant species within the continental U.S. For this analysis, we developed a ranking of predictors *a priori* to guide predictor selection for each species based on its natural history, such as winter annual species which use overwinter and spring moisture. We first grouped predictors into ten broad categories and ranked those categories based on our experience developing models for > 140 invasive plants in the continental U.S. (Young et al. 2020; Engelstad et al. 2022) and what environmental characteristics are important for different plant life forms in general. Next, we ranked the predictors within each of these broad categories for each species based on natural history knowledge of each individual species. Beginning with the highest ranked category (ETo), the highest ranked predictor was selected. Then, in the second ranked category, the highest predictor was selected that was not correlated with the first selected predictor for the top ranked category (maximum correlation coefficient of Pearson, Spearman, or Kendall was > 0.7 (Dormann et al. 2013)). An exception was made such that if one category of predictors was eliminated entirely, the second ranked predictor in category 1 would be retained if the highest was correlated with all of category 2 predictors whereas the 2nd ranked allowed for inclusion of another predictor category. Thus, correlation coefficients among predictors were used to limit co-linearity of selected predictors, but correlations with the response variable were not considered in variable selection. We ensured the ratio of presence points to predictors was at least 10:1 (Hosmer and Lemeshow 2000). This resulted in 47 predictors used across all models (Suppl. material 1: Table S4), with a range of 8 to 29 predictors per model.

Analyses

We evaluated the degree to which each species was disproportionately found within sagebrush and within different land cover types by overlaying occurrence points with land cover data. We identified where each focal species has invaded sagebrush communities by overlaying the compiled occurrence data with the NLCD shrubland sagebrush rangeland fractional component product (Xian et al. 2015; U.S. Geological Survey and Rigge 2019), defining sagebrush as any location with a > 0 cover value. We also counted the number of presence points for each species within the broad 2016 National Land Cover Data classes (i.e., agricultural, developed, forest, grassland, shrubland). We then calculated the proportion of the total species points found in each class. Because sampling effort can distort distributional assessments (e.g., Sofaer and Jarnevich 2017), we controlled for sampling effort across land cover categories by adjusting observed focal species associations by the habitat-specific number of records for other non-native plants of the same life form (i.e., forb/ herb or graminoid). We plotted the results to assess the degree to which each species disproportionately occurred in sagebrush habitats and each land cover class to better understand habitat preferences and the degree to which different species were sage specialists. These results were interpreted visually, while the target-background method described below similarly accounted for sampling biases within models.

We developed an ensemble of species distribution models for each species and training extent combination containing at least 50 presence locations (Suppl. material 1: Table S3). We fit models using the VisTrails Software for Assisted Habitat Modeling v2.2.0 (SAHM; Morisette et al. 2013) following the methods of Young et al. (2020) and high performance computing (Falgout and Gordon 2021). We implemented five model algorithms [boosted regression tree (Elith et al. 2008), generalized linear model (McCullagh and Nelder 1989), multivariate adaptive regression spline (Elith and Leathwick 2007), Maxent (Phillips et al. 2017), and random forest (Breiman 2001)] and two background point generation methods. One method was a kernel density estimate (KDE) around presence points to weight random background point generation (Elith et al. 2010). The other was target background (Phillips et al. 2009) based on 10,000 randomly selected locations of other non-native species locations within the same broad life form assigned by USDA Plants Database [forb/herb or graminoid] from within a 99% kernel density estimate isopleth (an isopleth is a line representing a constant value, as in a contour line on a topographical map) around the presence points or the full set of life form points if < 10,000 points fell within the 99% KDE. KDE isopleths are commonly used to define species' ranges by drawing a polygon to encompass locations (in this case, 99% of them) (Worton 1989) and recommended for range shifting invasive species (Elith et al. 2010). We withheld presence and background locations falling within test strips from estimation of all models. We fit each model using SAHM default parameters for algorithms with 10-fold cross-validation. We examined the difference between train and mean cross-validation values from the area under the receiver operating characteristic curve (AUC) and visually examined response curves to determine if models appeared overfit. In cases where models were

deemed overfit (trainAUC – testAUC > 0.05 or overly complex response curves), we adjusted model-specific tuning parameters, making the changes that most decreased overfitting while maintaining good cross-validation performance.

Because we only had presence locations, the outputs of the SDM algorithms are interpreted as relative habitat suitability values rather than probabilities. To create an ensemble across algorithms and background methods (10 models) we used of the 10th percentile training presence threshold for each model to produce binary outputs of suitable/unsuitable habitat that we could then sum across the ten models for each species/extent combination. The 10th percentile threshold is calculated for presence-only data based on the omission rate, where the 10% of occurrences with lowest predicted suitability are assumed to occur in poor habitat to avoid over-prediction due to errors or outliers in training locations.

We compared variable importance between regional and national models. We calculated variable importance by permutating values for each predictor across presence and background locations and calculating the difference between the original and permutated AUC values. Within each model, variables were ranked by permutation importance, with the most important variable being the one for which permuting its values led to the greatest decrease in AUC. For the ensemble we averaged the importance across the contributing models.

Because AUC is problematic for presence-background data (Lobo et al. 2008; Peterson et al. 2008; Sofaer et al. 2019a; Jiménez and Soberón 2020), we used the Boyce index to evaluate model performance based on the test data (Hirzel et al. 2006). The Boyce index assesses how much model predictions differ from random expectations by comparing the proportion of occurrences across classes of predicted suitability to the proportion of grid cells in each class. The Boyce index is based on the null expectation that the proportion of validation points expected within a given class is the proportion of the landscape area within that class (i.e., in contrast to sensitivity, which is based only on true positives, it would penalize a model for predicting high suitability everywhere). We calculated the index using the ensemble value (the number of models predicting suitable habitat) as the class bin for the Boyce index, generating 11 classes corresponding to the ensemble values of 0 to 10. Moving from low (zero models predicting suitability) to high (all 10 models predicting suitability), a high performing model will have a higher density of validation points at high ensemble values. Thus, the Boyce index is the Spearman rank correlation between the ordered classes (0–10 in our case) and the proportion of validation points in the focal class divided by the proportion of area in that class. We restricted validation points (Suppl. material 1: Table S3) and area calculations to sagebrush areas within each test strip (Fig. 1). We compared the Boyce index between national and regional training extents for each species, such that each model ensemble was tested on the same set of points for a given species and region.

We also compared the area within our three focal regions predicted to be suitable by each model ensemble. To do this, we turned the ensemble maps into binary suitable/ unsuitable maps by classifying any pixel within the region with an ensemble value of 6 or greater as suitable. We then counted the number of suitable pixels anywhere within each of the three different regions for each model ensemble.

Data accessibility statement

The data underpinning the analysis reported in this paper are available by a U.S. Geological Survey data release through the Science Base Repository at https://doi.org/10.5066/P90AL0PN.

Results

Most of our focal invasive plants had higher proportions of occurrences in sagebrush habitats compared to occurrences of all invasive plants of the same life form, pointing towards preference for sagebrush habitats after accounting for potential variation in sampling intensity with habitat type. *Ventenata dubia* occurred in sagebrush habitats in a greater proportion relative to occurrence points of other graminoid invasive species, as did *T. caput-medusae* to a lesser extent (Suppl. material 1: Fig. S2). Of the forb/herb species, *C. virgata*, *C. juncea*, and *S. aethiopis* also had positive ratios for sagebrush, indicating that these five species are disproportionately problematic within sagebrush habitats, even after considering sampling biases in occurrence locations. The three *Centaurea* species, *C. juncea*, and *H. glomeratus* all had positive ratios for the eastern sage region compared to other invasive forb species. All species had a positive association with shrubland, which includes sagebrush dominated locations, except *C. juncea* which had a positive association with the herbaceous land cover classes (Suppl. material 1: Fig. S3). It did, however, still have a positive ratio of occurrences in sagebrush everywhere but the Great Plains region (Suppl. material 1: Fig. S2).

Only two species, *C. diffusa* and *R. repens*, had enough locations in all three regions to fit models to all model estimation extents (Suppl. material 1: Table S3, Fig. S1). Patterns in predictions between the different training extents varied by species. *R. repens* mapped predictions varied with the training extent (Fig. 2), but the total area within each region predicted to be suitable by each model ensemble varied less for *R. repens* than for *C. diffusa* (*R. repens* points were closer to 1:1 line in Fig. 3b; Suppl. material 1: Fig. S4a, g). *Centauria diffusa* model ensembles that were trained on occurrences in sage showed poor extrapolation to other habitats in that they were less restricted to sage compared to the national model ensemble (i.e., sage only models, represented by red triangles in Fig. 3b, fell above the 1:1 line in Fig. 3b); interestingly, several of the national model ensembles for this species predicted less suitable area than their regional counterparts. Variable patterns could be seen for other species, with no clear visual differences in the geographic extent of predicted suitability among models trained on different regions (Suppl. material 1: Fig. S4). Some regional models predicted a lot of suitable habitat outside their training region, potentially extrapolating incorrectly (Suppl. material 1: Fig. S4); models extrapolated to other regions could show higher or lower suitable area than continental models, with extrapolation leading to more variability than interpolation (i.e., the points farthest from the 1:1 line in Fig. 3b are small, indicating they arose via extrapolation). Important predictors between training extents were relatively similar (Suppl. material 1: Fig. S5).
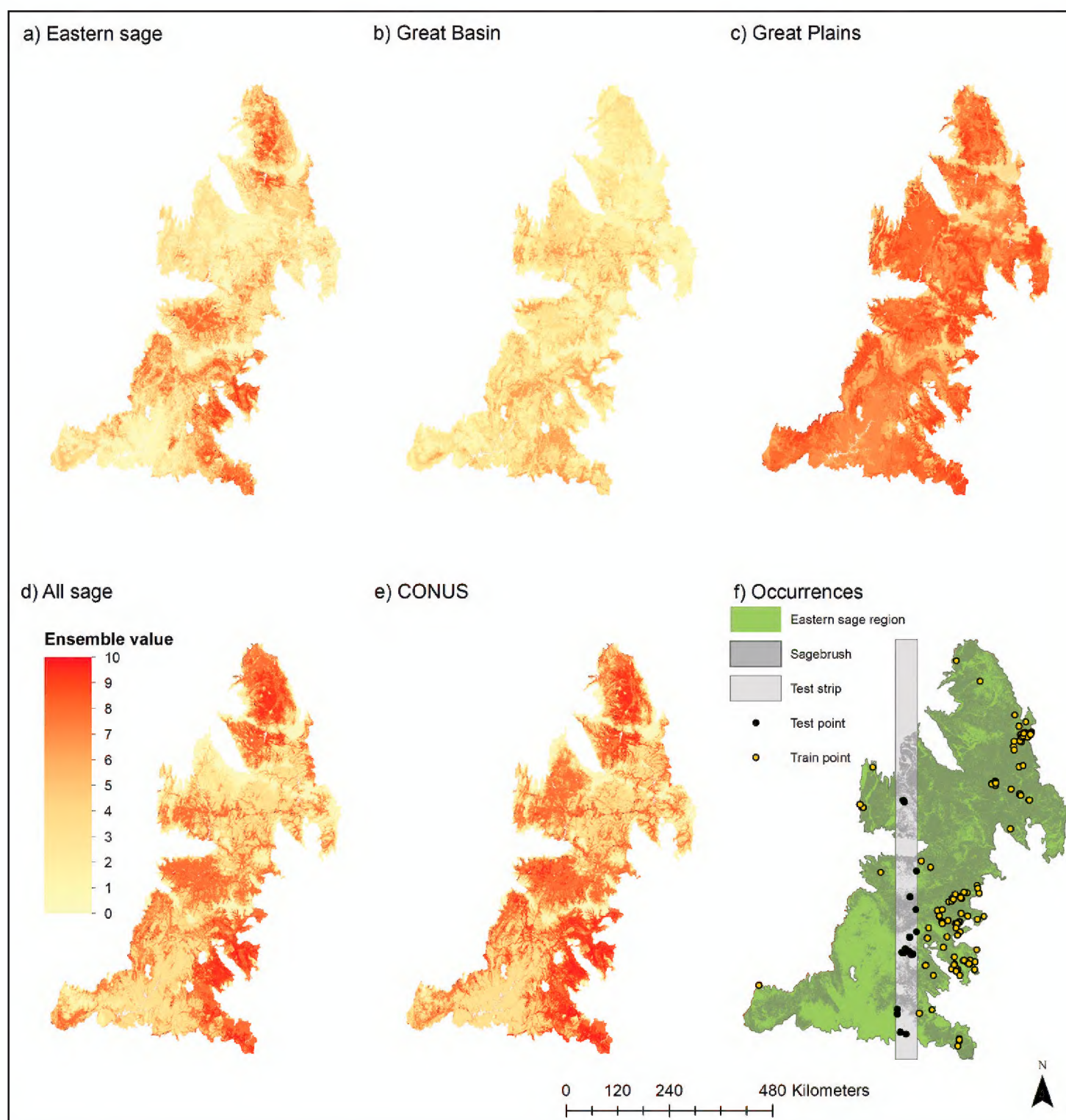
Figure 2. Predicted suitability for *Rhaponticum repens* within the eastern sage region (green region in Fig. 1) from models trained using data from the **A** Eastern sage **B** Great Basin **C** Great Plains **D** All sage, and **E** the continental U.S. (CONUS) (training regions shown in Fig. 1). Maps (**A–E**) show ensembled model predictions, defined as the number of models predicting suitable habitat; **F** shows training and test data for *Rhaponticum repens* within Eastern sage; test data were withheld from estimation of all models and used to create consistent performance assessment sets for each species and region.

Models tested on the region where they were trained were not better than continental U.S. models (paired t-test p-value = 0.07, mean difference = -0.14, i.e., continental models marginally better). Continental U.S. models outperformed models trained on the test region in seven of ten cases (Fig. 3). We had 11 regional test datasets across the nine species which met our criterion of 30 test points within the withheld spatial strip (Suppl. material 1: Table S3). Of these, the continental models or all sagebrush models were better than regional models (including those trained in other
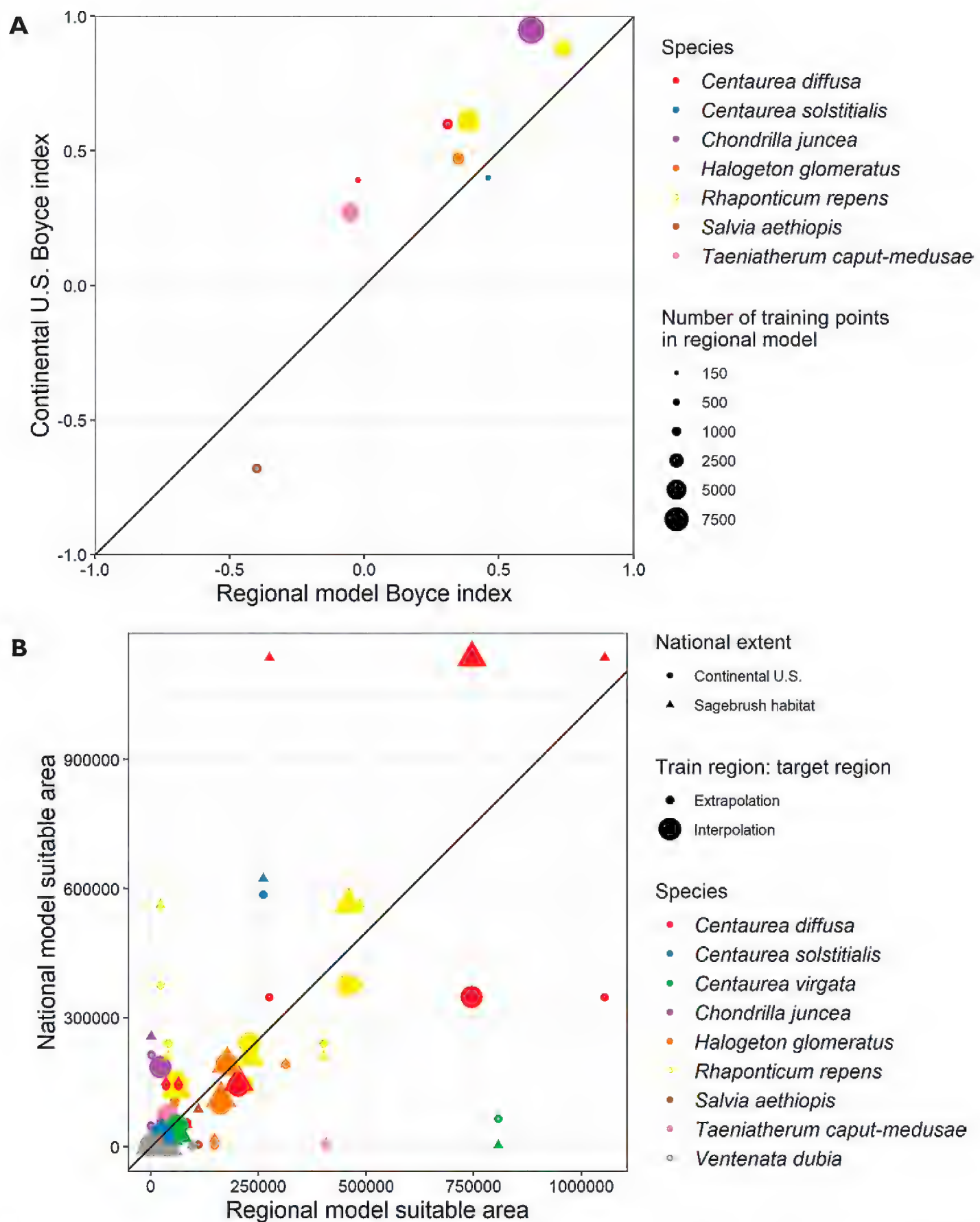
Figure 3. A Regional models did not outperform continental-scale models, even when many points were available within the training region. Boyce index values were calculated for the training region's test strip for both the matching region model ensemble (x-axis) and continental United States model ensemble (y-axis) for each species (color). Species without sufficient occurrence points within the test strip were excluded. Values above the 1-1 line indicate continental U.S. model had better performance; for most species and regions, models with a continental extent performed better even when the number of regional training points was high (i.e., points are above the 1:1 line, even for big points). **B** suitable area predicted by national models (either entire continental U.S. or sagebrush habitat within the U.S.) compared to regional models, where larger size indicated if the focal region considered for area calculation was the same (interpolation) or different (extrapolation) from the regional modeling training region. Values above 1-1 line indicate the national model predicted more suitable habitat.

regions) in most cases (Fig. 3, Suppl. material 1: Fig. S6). A regional model performed better than any other in five cases, of which three were actually models trained in other regions and extrapolated to the test region. The two models trained on the same region as the test strip had a lower Boyce index than models trained on either all sage or other regions. All models for *S. aethiopis* were poor, with all Boyce index values well below zero, despite decent performance according to typical assessment metrics based on the training data (cross validation AUC > 0.75, with an average per region > 0.89).

While *V. dubia* had enough locations to meet our criteria to develop models for the Great Plains region (n = 4,246), the occurrences were all within a relatively small geographic extent, and there were not enough locations for validation (Suppl. material 1: Table S3, Fig. S1). This small geographic extent was problematic in fitting models, where we were unable to obtain enough target background locations within the area around the general extent of occurrences within the region. Three of the five KDE models had poor fit (e.g., training AUC = 0.67 (GLM), 0.695 (MARS), 0.64 (RF)) that we were unable to improve; the other two KDE models had moderate performance (AUC < 0.8).

Discussion

Regionally trained models for invasive plants of management concern did not perform better than national models when evaluated with independent data from within the training region. Continental-scale models tended to outperform regional ones even when the number of regional training points was high (Fig. 3), supporting the general recommendation to use a broad spatial extent for training models of invasive species (Fitzpatrick et al. 2007; Broennimann and Guisan 2008). Mapped predictions from models trained on a focal region were more similar to continental scale predictions within that region, compared to extrapolated results from models trained in other regions (Fig. 2; Suppl. material 1: Fig. S4). When comparing area predicted as suitable by models trained on different geographic extents for the same target region, there was not a consistent pattern, but extrapolation led to more variable results (Fig. 3b). When interpolating, including training points beyond the focal region did affect predictions within that region, as we found differences in both the spatial pattern and the overall level of predicted suitability between continental and regional model outputs. The tendency for higher performance of continental models points to these modifications being generally positive for within-region model performance and indicates that models with a broader extent are less prone to swamp regional patterns than to usefully reduce model extrapolation.

For most species, we had insufficient data to estimate and evaluate a model in one or more of our focal sagebrush regions. For example, *V. dubia* lacked estimation data in the eastern sage region, and is established within only a small area of the Great Plains, where active EDRR efforts have yielded a large number of data points (Hart and Mealor 2021). However, because our validation design utilized spatial strips to provide a more independent, and therefore more realistic, estimate of performance, we had insufficient regional validation points to assess model performance within the

Great Plains. In addition, strong spatial clustering of points early in an invasion, such as with *V. dubia* in the Great Plains, can reflect propagule pressure and the idiosyncrasies of dispersal, with many unoccupied locations due to dispersal limitation (Elith et al. 2010; Václavík and Meentemeyer 2012). Species distribution models trained on only a portion of a species' range are therefore likely to be less accurate in early invasion stages.

While this study focused on the geographic extent of estimation data, comparisons with previous work highlight how other modeling decisions shape predicted invasion risk. Here, we thresholded individual models in our ensembles based on a rule that categorized 10% of training presences as occurring in unsuitable habitat. This threshold rule is appropriate for EDRR activities where search is the end use of models and a targeted approach can focus search efforts towards areas with a relatively higher degree of suitability (Sofaer et al. 2019b). In contrast, Jarnevich et al. (2021) quantified invasion risk across management units, and therefore used a more precautionary approach, the 1st percentile threshold, to avoid minimizing invasion risk via errors of omission. In contrast to the 10th percentile threshold, the first percentile classifies 1% of training presences as being in unsuitable habitat and thus classifies a larger portion of a study area as suitable. Both thresholds are based only on presence information, as true absences are unavailable. The more targeted threshold used here resulted in a smaller extent of predicted suitability for the same species, and illustrates how different thresholds may be implemented depending on intended use (Sofaer et al. 2019b).

Our study varied the geographic extent of estimation data to compare continental and regional models. Our findings align with results for native species, where in the absence of *a priori* evidence for niche divergence, researchers recommended creating models across a species' range (Collart et al. 2021; Connor et al. 2019). However, we held predictor variables constant between geographies, and the inclusion of geospatial variables believed important for controlling a species distribution may produce a better model than one for a larger extent lacking that information. Indeed, our continental models do not include species' global ranges because we highly value predictor variables that are available for the U.S. but are not available, inconsistent, or of lower quality globally (e.g., information on soils). For these species we lacked information that would indicate we needed to vary predictors geographically.

Alternatives to regional models include allowing for non-stationarity in environmental responses via hierarchical modeling, geographically weighted regression (Osborne et al. 2007) or spatially-varying coefficient models (Gelfand et al. 2003). Hierarchical models can estimate both overall environmental responses and variation in those relationships among groups (e.g., via random slopes in a mixed modeling framework). Both regional and hierarchical modeling approaches typically require defining intraspecific groups, but little emphasis has been placed on the approaches used to define subpopulations, which should be well justified (Chardon et al. 2020). Here, we considered intraspecific divisions based on ecoregions; among native species, studies have diverged in whether splitting by ecoregion (Smith et al. 2019a) or by genetic similarity (Chardon et al. 2020) yields the best performance. Partial pooling, a hierarchical approach that incorporates multiple intraspecific groups within a single mixed

model, provides a method intermediate between splitting and lumping (Smith et al. 2019b). The research to define subpopulations takes time and resources which may not be available for many invasive species, particularly when time to action is critical in limiting invasion costs (Pergl et al. 2020). These alternatives add complexity and potentially require more resources to first define groupings appropriately and then create multiple or hierarchical models for a single taxon compared to a continental approach. There is a continuum of automation versus human time and insight in developing species distribution models (Young et al. 2020), from large extent models for suites of species using the same predictors for all models (e.g., Allen and Bradley 2016) to very detailed models for a single species (e.g., Smith et al. 2019b; Chardon et al. 2020). The best path forward depends on the objectives, data availability, a priori information about populations and species, and the available resources and timeline.

In selecting a modeling approach, it is important to distinguish between populations that have different limiting factors and populations that have different responses to environmental conditions. Across a species' range, it is typical that different limiting factors are suspected to constrain population growth; for example, an early macroecological hypothesis posited that biotic interactions more often defined southern range limits while abiotic conditions more often defined northern range limits (reviewed by Schemske et al. 2009). Cases where, for example, one area may be too dry while another is too cold can be handled by standard range-wide modeling approaches, as demonstrated by our study. It is only where the definition of 'too cold' varies among populations that more tailored or complex models may be warranted as highlighted by other studies of native species. Ideally, common garden experiments and similar tools would be used to test for differentiation but these types of studies for every invasive species would be time and cost prohibitive.

Conclusion

The degree of variation in responses to environmental conditions and the amount of data available underlie the selection of appropriate strategies for species distribution modeling (Fig. 4). Consistent responses to ecological conditions (e.g., Connor et al. 2019; Collart et al. 2021) or capturing a broader range of environmental conditions occupied by a species (e.g., Broennimann and Guisan 2008) support range-wide modeling (bottom right), while evidence for regional differentiation lends support to regional or hierarchical modeling methods where data are available (e.g., Chardon et al. 2020 upper right). However, there is a key tension between data availability and relevance for EDRR. Model outputs inform EDRR when they can be used to guide surveillance efforts and assess spatial patterns of invasion risk during a rapid response. Yet at these early stages of an invasion, there is necessarily little to no data on species' occurrences within the focal area or the data occur within such a small extent that model fitting is difficult (e.g., *V. dubia* in the Great Plains region; left side of Fig. 4). Regional models will typically be most relevant at later stages of an invasion, where
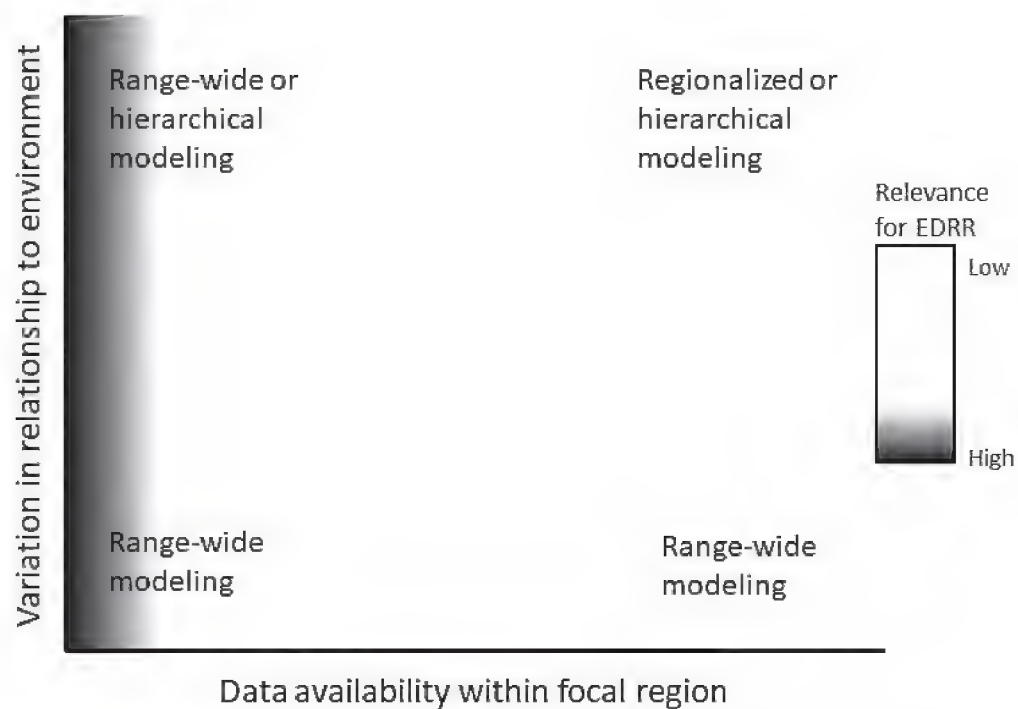
Figure 4. Conceptual depiction of the utility of different modeling methods and of the trade-offs between data availability within a focal region and relevance of model outputs for Early Detection and Rapid Response (EDRR) within that region. Range-wide modeling is appropriate where there is little variation in the relationship between a species' occurrence and environmental conditions. Where local populations are differentiated in their responses to the environment, hierarchical or regionalized models are expected to produce the most relevant predictions for within the region, and the selection among model types may depend on data availability, institutional capacity, and time horizon for delivering results. The relevance of model outputs for EDRR is high only very early in an invasion, when few data are available; therefore, range-wide modeling is expected to remain the primary tool used to anticipate habitat suitability for non-native species.

there has been more opportunity for population divergence, range filling, and data collection (moving from left to right within Fig. 4). Clear justification and communication of model assumptions between model producers, local knowledge holders, and decision-makers can help clarify what kinds of differences warrant regional or hierarchical models. Delayed actions may increase costs associated with invasions and decrease the ability to meet management goals for newly introduced species to a region (Ahmed et al. 2022). Regional models did not perform better than national models, and thus national models may have use to inform management decisions for early detection of invasive species.

Acknowledgements

References

Ahmed DA, Hudgins EJ, Cuthbert RN, Kourantidou M, Diagne C, Haubrock PJ, Leung B, Liu C, Leroy B, Petrovskii S, Beidas A, Courchamp F (2022) Managing biological invasions: The cost of inaction. Biological Invasions 24(7): 1927–1946. https://doi.org/10.1007/s10530-022-02755-0

Allen JM, Bradley BA (2016) Out of the weeds? Reduced plant invasion risk with climate change in the continental United States. Biological Conservation 203: 306–312. https://doi.org/10.1016/j.biocon.2016.09.015

Barbet-Massin MA-OX, Rome Q, Villemant C, Courchamp F (2018) Can species distribution models really predict the expansion of invasive species? PLoS ONE 13(3): e0193085. https://doi.org/10.1371/journal.pone.0193085

Breiman L (2001) Random forests. Machine Learning 45(1): 5–32. https://doi.org/10.1023/A:1010933404324

Broennimann O, Guisan A (2008) Predicting current and future biological invasions: Both native and invaded ranges matter. Biology Letters 4(5): 585–589. https://doi.org/10.1098/rsbl.2008.0254

Brooks ML, Klinger RC (2009) Practical considerations for early detection monitoring of plant invasions. In: Inderjit (Ed.) Management of invasive weeds. Springer, Heidelberg, 9–33. https://doi.org/10.1007/978-1-4020-9202-2_2

Chamberlain S, Szocs E (2013) "taxize - taxonomic search and retrieval in R." F1000Research. https://doi.org/10.12688/f1000research.2-191.v2

Chamberlain S, Szoecs E, Foster Z, Arendsee Z, Boettiger C, Ram K, Bartomeus I, Baumgartner J, O'Donnell J, Oksanen J, Tzovaras B, Marchand P, Tran V, Salmon M, Li G, Grenié M (2020) taxize: Taxonomic information from around the web. R package version 0.9.98. https://github.com/ropensci/taxize

Chardon NI, Pironon S, Peterson ML, Doak DF (2020) Incorporating intraspecific variation into species distribution models improves distribution predictions, but cannot predict species traits for a wide-spread plant species. Ecography 43(1): 60–74. https://doi.org/10.1111/ecog.04630

Coates PS, Ricca MA, Prochazka BG, Brooks ML, Doherty KE, Kroger T, Blomberg EJ, Hagen CA, Casazza ML (2016) Wildfire, climate, and invasive grass interactions negatively impact an indicator species by reshaping sagebrush ecosystems. Proceedings of the National Academy of Sciences of the United States of America 113(45): 12745–12750. https://doi.org/10.1073/pnas.1606898113

Colautti RI, Barrett SCH (2013) Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant. Science 342(6156): 364–366. https://doi.org/10.1126/science.1242121

Collart F, Hedenäs L, Broennimann O, Guisan A, Vanderpoorten A (2021) Intraspecific differentiation: Implications for niche and distribution modelling. Journal of Biogeography 48(2): 415–426. https://doi.org/10.1111/jbi.14009

Connor T, Viña A, Winkler JA, Hull V, Tang Y, Shortridge A, Yang H, Zhao Z, Wang F, Zhang J, Zhang Z, Zhou C, Bai W, Liu J (2019) Interactive spatial scale effects on species distribution modeling: The case of the giant panda. Scientific Reports 9(1): e14563. https://doi.org/10.1038/s41598-019-50953-z

Crist MR, Chambers JC, Phillips SL, Prentice KL, Wiechman LA [Eds] (2019) Science framework for conservation and restoration of the sagebrush biome: Linking the Department of the Interior's Integrated Rangeland Fire Management Strategy to long-term strategic conservation actions. Part 2. Management applications. Gen. Tech. Rep. RMRS-GTR-389. U.S Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 237 pp. https://doi.org/10.2737/RMRS-GTR-389

Davies KW, Boyd CS, Beck JL, Bates JD, Svejcar TJ, Gregg MA (2011) Saving the sagebrush sea: An ecosystem conservation plan for big sagebrush plant communities. Biological Conservation 144(11): 2573–2584. https://doi.org/10.1016/j.biocon.2011.07.016

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 027–046. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Elith J, Leathwick J (2007) Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. Diversity & Distributions 13(3): 265–275. https://doi.org/10.1111/j.1472-4642.2007.00340.x

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. Journal of Animal Ecology 77(4): 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods in Ecology and Evolution 1(4): 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Engelstad P, Jarnevich CS, Hogan T, Sofaer HR, Pearse IS, Sieracki JL, Frakes N, Sullivan J, Young NE, Prevéy JS, Belamaric P, LaRoe J (2022) INHABIT: A Web-Based Decision Support Tool for Invasive Species Habitat Visualization and Assessment Across the Contiguous United States. PLoS ONE 17(2): e0263056. https://doi.org/10.1371/journal.pone.0263056

Falgout JT, Gordon J (2021) USGS Advanced Research Computing, USGS Yeti Supercomputer. US Geological Survey. https://doi.org/10.5066/F7D798MJ

Fitzpatrick MC, Weltzin JF, Sanders NJ, Dunn RR (2007) The biogeography of prediction error: Why does the introduced range of the fire ant over-predict its native range? Global Ecology and Biogeography 16(1): 24–33. https://doi.org/10.1111/j.1466-8238.2006.00258.x

Fourcade Y, Besnard AG, Secondi J (2018) Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Global Ecology and Biogeography 27(2): 245–256. https://doi.org/10.1111/geb.12684

Franklin J (2010) Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, New York, 338 pp. https://doi.org/10.1017/CBO9780511810602

GBIF.org (2021) GBIF Occurrence Download.

GBIF.org (2022) GBIF Occurrence Download. [Accessed 06 May 2022]

Gelfand AE, Kim H-J, Sirmans CF, Banerjee S (2003) Spatial Modeling With Spatially Varying Coefficient Processes. Journal of the American Statistical Association 98(462): 387–396. https://doi.org/10.1198/016214503000170

Hällfors MH, Liao J, Dzurisin J, Grundel R, Hyvärinen M, Towle K, Wu GC, Hellmann JJ (2016) Addressing potential local adaptation in species distribution models: Implications

for conservation under climate change. Ecological Applications 26(4): 1154–1169. https://doi.org/10.1890/15-0926

Hart M, Mealor BA (2021) Effects of Ventenata dubia removal on rangelands of northeast Wyoming. Invasive Plant Science and Management 14(3): 156–163. https://doi.org/10.1017/inp.2021.20

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199(2): 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hosmer DW, Lemeshow S (2000) Applied logistic regression. Wiley, New York, 373 pp. https://doi.org/10.1002/0471722146

Jarnevich CS, Talbert M, Morisette J, Aldridge C, Brown CS, Kumar S, Manier D, Talbert C, Holcombe T (2017) Minimizing effects of methodological decisions on interpretation and prediction in species distribution studies: An example with background selection. Ecological Modelling 363: 48–56. https://doi.org/10.1016/j.ecolmodel.2017.08.017

Jarnevich CS, Sofaer HR, Engelstad P (2021) Modelling presence versus abundance for invasive species risk assessment. Diversity & Distributions 27(12): 2454–2464. https://doi.org/10.1111/ddi.13414

Jarnevich CJ, Sofaer HR, Belamaric P, Engelstad P (2022) Data to fit and test models for invasive species in different geographic regions. U.S. Geological Survey data release. https://doi.org/10.5066/P90AL0PN

Jiménez L, Soberón J (2020) Leaving the area under the receiving operating characteristic curve behind: An evaluation method for species distribution modelling applications based on presence-only data. Methods in Ecology and Evolution 11(12): 1571–1586. https://doi.org/10.1111/2041-210X.13479

Liu C, Wolter C, Xian W, Jeschke JM (2020a) Most invasive species largely conserve their climatic niche. Proceedings of the National Academy of Sciences of the United States of America 117(38): 23643–23651. https://doi.org/10.1073/pnas.2004289117

Liu C, Wolter C, Xian W, Jeschke JM (2020b) Species distribution models have limited spatial transferability for invasive species. Ecology Letters 23(11): 1682–1692. https://doi.org/10.1111/ele.13577

Lobo JM, Jimenez-Valverde A, Real R (2008) AUC: A misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17(2): 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

McCullagh P, Nelder JA (1989) Generalized linear models, 2nd Edn. Chapman and Hall, London; New York. https://doi.org/10.1007/978-1-4899-3242-6

Morisette JT, Jarnevich CS, Holcombe TR, Talbert CB, Ignizio D, Talbert MK, Silva C, Koop D, Swanson A, Young NE (2013) VisTrails SAHM: Visualization and workflow management for species habitat modeling. Ecography 36(2): 129–135. https://doi.org/10.1111/j.1600-0587.2012.07815.x

Osborne PE, Foody GM, Suárez-Seoane S (2007) Non-stationarity and local approaches to modelling the distributions of wildlife. Diversity & Distributions 13(3): 313–323. https://doi.org/10.1111/j.1472-4642.2007.00344.x

Pearman PB, D'Amen M, Graham CH, Thuiller W, Zimmermann NE (2010) Within-taxon niche structure: Niche conservatism, divergence and predicted effects of climate change. Ecography 33(6): 990–1003. https://doi.org/10.1111/j.1600-0587.2010.06443.x

Pergl J, Pyšek P, Essl F, Jeschke JM, Courchamp F, Geist J, Hejda M, Kowarik I, Mill A, Musseau C, Pipek P, Saul W-C, von Schmalensee M, Strayer D (2020) Need for routine tracking of biological invasions. Conservation Biology 34(5): 1311–1314. https://doi.org/10.1111/cobi.13445

Peterson AT, Papes M, Soberon J (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecological Modelling 213(1): 63–72. https://doi.org/10.1016/j.ecolmodel.2007.11.008

Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecological Applications 19(1): 181–197. https://doi.org/10.1890/07-2153.1

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME (2017) Opening the black box: An open-source release of Maxent. Ecography 40(7): 887–893. https://doi.org/10.1111/ecog.03049

Reaser JK, Burgiel SW, Kirkey J, Brantley KA, Veatch SD, Burgos-Rodríguez J (2020) The early detection of and rapid response (EDRR) to invasive species: A conceptual framework and federal capacities assessment. Biological Invasions 22(1): 1–19. https://doi.org/10.1007/s10530-019-02156-w

Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F, Dormann CF (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40(8): 913–929. https://doi.org/10.1111/ecog.02881

Schemske DW, Mittelbach GG, Cornell HV, Sobel JM, Roy K (2009) Is There a Latitudinal Gradient in the Importance of Biotic Interactions? Annual Review of Ecology, Evolution, and Systematics 40(1): 245–269. https://doi.org/10.1146/annurev.ecolsys.39.110707.173430

Sequeira AMM, Bouchet PJ, Yates KL, Mengersen K, Caley MJ (2018) Transferring biodiversity models for conservation: Opportunities and challenges. Methods in Ecology and Evolution 9(5): 1250–1264. https://doi.org/10.1111/2041-210X.12998

Smith AB, Beever EA, Kessler AE, Johnston AN, Ray C, Epps CW, Lanier HC, Klinger RC, Rodhouse TJ, Varner J, Perrine JD, Seglund A, Hall LE, Galbreath K, MacGlover C, Billman P, Blatz G, Brewer J, Castillo Vardaro J, Chalfoun AD, Collins G, Craighead A, Curlis C, Daly C, Doak DF, East M, Edwards M, Erb L, Ernest KA, Fauver B, Foresman KR, Goehring K, Hagar J, Hayes CL, Henry P, Hersey K, Hilty SL, Jacobson J, Jeffress MR, Manning T, Masching A, Maxell B, McCollough R, McFarland C, Miskow E, Morelli TL, Moyer-Horner L, Mueller M, Nugent M, Pratt B, Rasmussen-Flores M, Rickman TH, Robison H, Rodriguez A, Rowe K, Rowe K, Russello MA, Saab V, Schmidt A, Stewart JAE, Stuart JN, Svancara LK, Thompson W, Timmins J, Treinish G, Waterhouse MD, Westover ML, Wilkening J, Yandow L (2019a) Alternatives to genetic affinity as a context for within-species response to climate. Nature Climate Change 9(10): 787–794. https://doi.org/10.1038/s41558-019-0584-8

Smith AB, Godsoe W, Rodríguez-Sánchez F, Wang H-H, Warren D (2019b) Niche Estimation Above and Below the Species Level. Trends in Ecology & Evolution 34(3): 260–273. https://doi.org/10.1016/j.tree.2018.10.012

Sofaer HR, Jarnevich CS (2017) Accounting for sampling patterns reverses the relative importance of trade and climate for the global sharing of exotic plants. Global Ecology and Biogeography 26(6): 669–678. https://doi.org/10.1111/geb.12577

Sofaer HR, Hoeting JA, Jarnevich CS (2019a) The area under the precision-recall curve as a performance metric for rare binary events. Methods in Ecology and Evolution 10(4): 565–577. https://doi.org/10.1111/2041-210X.13140

Sofaer HR, Jarnevich CS, Pearse IS, Smyth RL, Auer S, Cook GL, Edwards Jr TC, Guala GF, Howard TG, Morisette JT, Hamilton H (2019b) Development and Delivery of Species Distribution Models to Inform Decision-Making. Bioscience 69(7): 544–557. https://doi.org/10.1093/biosci/biz045

U.S. Environmental Protection Agency (2013) Level III Ecoregions of the Conterminous United States. U.S. EPA Office of Research and Development (ORD) - National Health and Environmental Effects Research Laboratory (NHEERL), Corvallis, OR.

U.S. Geological Survey, Rigge M (2019) National Land Cover Database (NLCD) 2016 Shrubland Fractional Components for the Western U.S. (ver. 1.0, May 2019): U.S. Geological Survey data release.

USDA NRCS (2019) The PLANTS Database (http://plants.usda.gov, 05/13/2019). National Plant Data Team, Greensboro, NC USA.

Václavík T, Meentemeyer RK (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Diversity & Distributions 18(1): 73–83. https://doi.org/10.1111/j.1472-4642.2011.00854.x

Vale CG, Tarroso P, Brito JC (2014) Predicting species distribution at range margins: Testing the effects of study area extent, resolution and threshold selection in the Sahara–Sahel transition zone. Diversity & Distributions 20(1): 20–33. https://doi.org/10.1111/ddi.12115

Werkowska W, Márquez AL, Real R, Acevedo P (2017) A practical overview of transferability in species distribution modelling. Environmental Reviews 25(1): 127–133. https://doi.org/10.1139/er-2016-0045

Worton BJ (1989) Kernel methods for estimating the utilization distribution in home-range studies. Ecology 70(1): 164–168. https://doi.org/10.2307/1938423

Xian G, Homer C, Rigge M, Shi H, Meyer D (2015) Characterization of shrubland ecosystem components as continuous fields in the northwest United States. Remote Sensing of Environment 168: 286–300. https://doi.org/10.1016/j.rse.2015.07.014

Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, Parnell S, Fielding AH, Bamford AJ, Ban S, Barbosa AM, Dormann CF, Elith J, Embling CB, Ervin GN, Fisher R, Gould S, Graf RF, Gregr EJ, Halpin PN, Heikkinen RK, Heinänen S, Jones AR, Krishnakumar PK, Lauria V, Lozano-Montes H, Mannocci L, Mellin C, Mesgaran MB, Moreno-Amat E, Mormede S, Novaczek E, Oppel S, Ortuño Crespo G, Peterson AT, Rapacciuolo G, Roberts JJ, Ross RE, Scales KL, Schoeman D, Snelgrove P, Sundblad G, Thuiller W, Torres LG, Verbruggen H, Wang L, Wenger S, Whittingham MJ, Zharikov Y, Zurell D, Sequeira AMM (2018) Outstanding Challenges in the Transferability of Ecological

Models. Trends in Ecology & Evolution 33(10): 790–802. https://doi.org/10.1016/j.tree.2018.08.001

Young NE, Jarnevich CS, Sofaer HR, Pearse I, Sullivan J, Engelstad P, Stohlgren TJ (2020) A modeling workflow that balances automation and human intervention to inform invasive plant management decisions at multiple spatial scales. PLoS ONE 15(3): e0229253. https://doi.org/10.1371/journal.pone.0229253

Supplementary material I

Tables S1–S4, Figures S1–S6

Authors: Catherine S. Jarnevich, Helen R. Sofaer, Pairsa Belamaric, Peder Engelstad

Data type: Supplemental figures and tables.

Explanation note: **Table S1.** Total area (km2) and area of sagebrush habitat within each modeling region and its associated test strip. **Table S2.** Model assessment rubric from Sofaer et al. (2019b) for presence models for all species modeled. **Table S3.** Location summary including number of training points (and test strip points) in each spatial extent considered including the Great Basin, eastern sagebrush, the Great Plains where all three are limited to sagebrush, sagebrush in the western United States, and all locations in the continental United States. **Table S4.** The predictors considered in model development including the variable (name [unit of measure | spatial (cell) resolution | temporal resolution (if applicable)]), description, and source from Engelstad et al. (2022). **Figure S1.** For each species, map showing location point data colored by region (Great Basin sagebrush, eastern sagebrush, Great Plains sagebrush, other sagebrush, and outside of sagebrush). **Figure S2.** Proportion of locations for each species found in sagebrush habitat, as defined by the National Land Cover Dataset shrubland sagebrush rangeland fractional component product (Xian et al. 2015), within each region (continental United States, eastern sage, the Great Basin, and the Great Plains). **Figure S3.** Proportion of the total locations for each species recorded in each of six National Land Cover Dataset categories relative to the proportion of all invasive species locations in the same life-form (forb/ herb for first seven species; graminoid for last two species) to correct for sampling biases. Values above 1 indicate that more locations are found in that land-cover class relative to all other species. **Figure S4.** The 10-percentile ensemble map (values of 0 to 10) for each modeled training area (Great Basin, eastern sagebrush, Great Plains, all sagebrush, and Continental U.S. with inclusion determined by Table S3), with output restricted to the three regions (Great Basin, eastern sagebrush, Great Plains) for each of the nine species (a to i). Each map is overlaid by a transparent gray layer highlighting novel environments from the Multivariate Environmental Similarity Surface. **Figure S5.** Ranked variable importance for each model run (algorithm by background method) across species plotted for models trained on regions (all three combined), all sagebrush, and the continental U.S. Difference in AUC values from permutation of occurrence and background locations were

ranked across predictors by training extent, species, algorithm, and background method. **Figure S6.** Boyce index calculated for each region's test strip (columns) and the test strip 10 percentile ensemble model by species (x-axis) including the number of the species' occurrences within the test strip above the axis for the different models including a model trained using species' locations from sage (all sage), the continental U.S., the region matching the test strip region (matching region), or regions different from the test strip region (other region).

Link: https://doi.org/10.3897/neobiota.77.86364.suppl1